Zander, R.H. 2014. Classical determination of monophyly, exemplified with *Didymodon* s. lat. (Bryophyta). Part 1 of 3, synopsis and simplified concepts. Phytoneuron 2014-78: 1–7. Published 6 August 2014. ISSN 2153 733X

CLASSICAL DETERMINATION OF MONOPHYLY, EXEMPLIFIED WITH *DIDYMODON* S. LAT. (BRYOPHYTA). PART 1 OF 3, SYNOPSIS AND SIMPLIFIED CONCEPTS

RICHARD H. ZANDER
Missouri Botanical Garden
P.O. Box 299
St. Louis, Missouri 63166
richard.zander@mobot.org

ABSTRACT

Classical systematists introduce evolutionary relationships into classification by identifying monophyly. The classical basis for determining monophyly as serial macroevolution of taxa is here put on a statistical basis. Clues to adaptive or relatively neutral transformative radiation are used in grouping ancestral species and descendant species. Such clues are additive in suggesting a direction of evolution of one taxon to another. These clues are usually intuitive to the informed expert but can be characterized in terms of probabilistic units, decibans. Since decibans are logarithmic they can be added to yield a probability of one species being ancestral to another. Transformative traits are convergent between groups, while conservative traits occur in multiple species and environments in one group. Stem-based evolutionary trees (caulograms) are generated by hypotheses of serial evolutionary change. Although the manner in which evolutionary relationships are evaluated is here considered standard in classical practice, the formalization of scientific heuristic intuition and method of documentation is new. It is hoped that this or a similar method of systematic analysis that creates classifications based on both serial and cladistic evolutionary relationships might be developed and expanded by innovative students and intellectually flexible researchers in systematics.

Given that this paper is mainly intended for motivated classical taxonomists and challengeable students, a synopsis of parts 2 and 3 is given here as a simplified guide to the fundamental concepts and methods. Part 2 gives technical justification and reasons for a new, combined classical-phylogenetic approach. It provides elements of the new method and its statistical basis. Part 3 is a morphologically based example of the method, showing serial evolutionary relationships among species of the moss genus *Didymodon* Hedw. and its close relatives (Pottiaceae, Bryophyta).

Some may feel uncomfortable that the here-recommended use of scientific inductive logic and discursive reasoning results in hypotheses of evolutionary relationships that are not as apparently well-supported as are cladograms in phylogenetics. The choice, however, is between (1) the best attempt at developing theory that addresses all facts including contradictory results in morphological and molecular studies (the method advocated in the present paper), and (2) an exact method that must be wrong much of the time (phylogenetics), as discussed below.

Justification

In morphological cladograms, some descendants from one ancestral species may duplicate traits (as parallelisms) of other descendants, creating false synapomorphies and false branches in a cladogram distal to the ancestral branch. This is probably fairly common given the few traits involved in speciation in any one genus. Or, if one or more descendant species has a reversal in a trait, such species appear lower in the cladogram than the ancestral branch. Examination of traits that are informative of serial transformations (not shared ancestry) can correct false resolution of branch order. A cladistic multifurcation (one ancestor generating multiple descendants) is then entirely appropriate. Evolution of species by serial transformation is here termed transformative radiation.

Parsimony analysis optimizes a cladogram by minimizing the number of trait transformations. Naming cladogram nodes further optimizes the cladogram by eliminating unnecessary unknown shared ancestors. Thus a "most-parsimonious" cladogram is made even more parsimonious by trimming off nodes that promise resolution that is not really based on shared ancestry (Fig. 1). This results in what is in effect a natural key (Zander 2013: 82), which can be directly represented by a caulogram. Naming cladogram nodes as extant taxa is referred to as "superoptimization."

A cladogram multifurcation remains a poor model of ancestor to multiple descendant relationships because the line below the node is the same taxon as one of the lines coming out of the multifurcation as in Fig. 1(2). The best diagram is a caulogram (or commagram, or Besseyan cactus). The polarity of evolutionary diagrams is standard with derived species towards the top, and ancestral species lower. Commonly a rooted, stand-alone cladogram or caulogram is understood to be a small area in the apex of a complex tree of life.

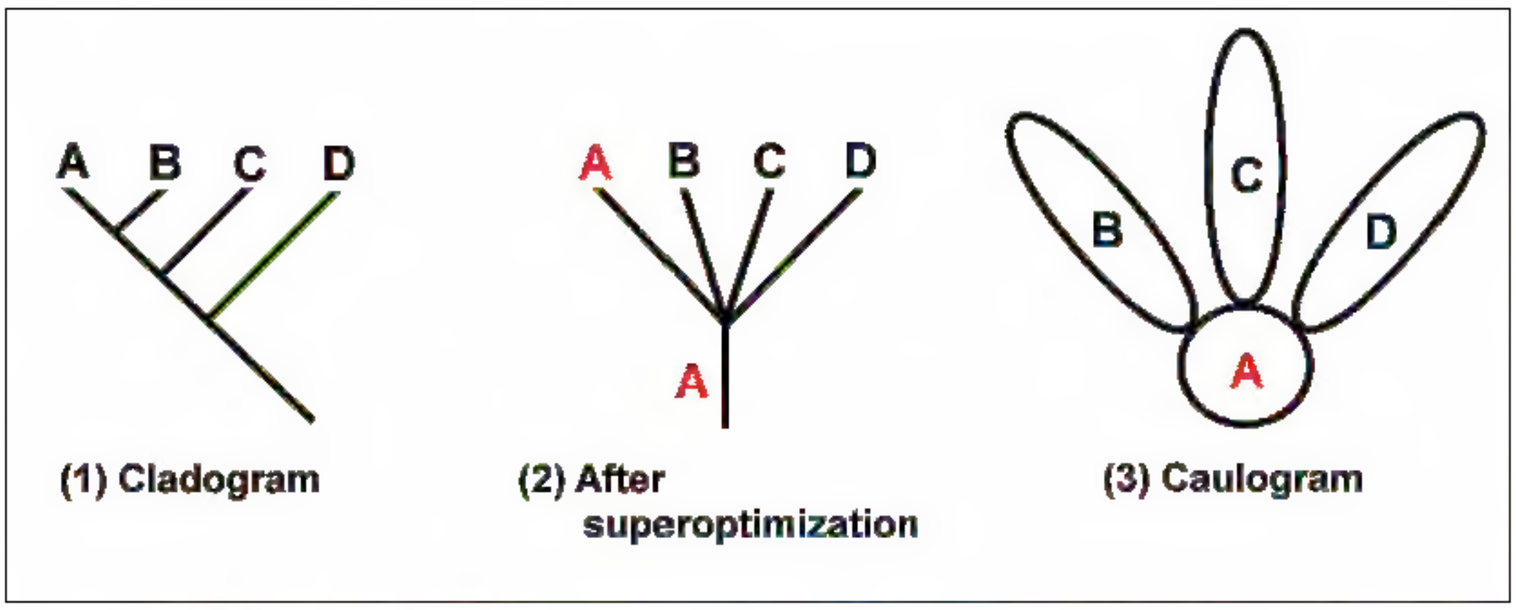

Figure 1. A cladogram (1), here a terminal branch of a larger rooted cladogram, can be changed to a more reasonable multifurcation (2) if A is found to be a taxon ancestral to B, C, and D after analysis for serial transformation. The node in the multifurcation (2) is identified as the same taxon as A. Because two cladogram branches both representing ancestor A in cladogram (2) are redundant, stem evolutionary relationships are better represented by a caulogram (3). This series simplifies how a cladogram is re-interpreted as a caulogram, wherein a tree of cladistic leaves becomes a tree of evolutionary stems.

In molecular studies, cladistic branch order of the molecular strains represented by specimens studied can be estimated to a large extent. However, broad-ranging ancestral species may have isolated molecular "strains" even though the molecular strains are essentially identical in expressed traits at the species level.

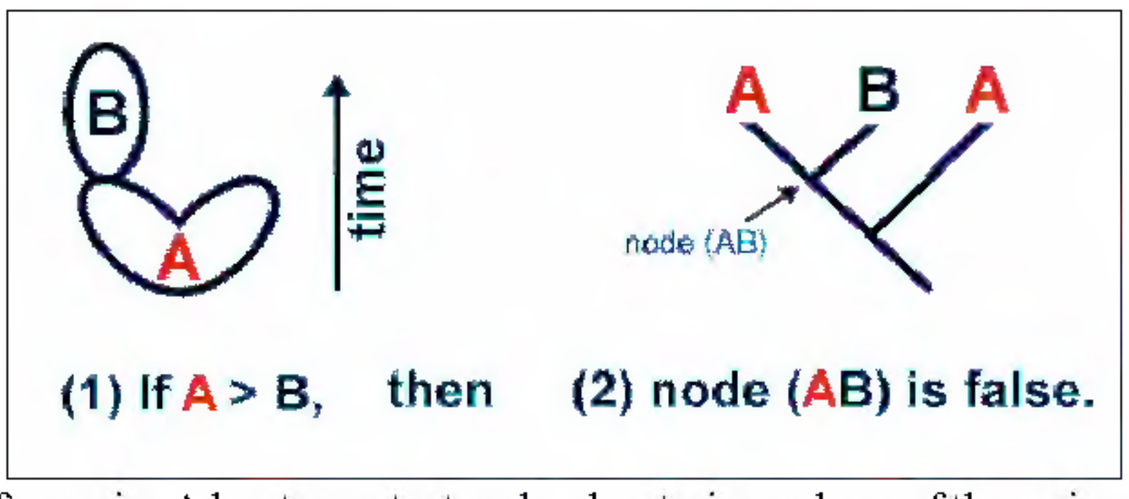

Figure 2. In (1) on left, species A has two extant molecular strains and one of them gives rise to (symbol ">") species B. Species A is then ancestral to species B. A molecular cladogram (2) distinguishes the molecular strains as separate with B arising from one of the strains. If species A is static in expressed traits (i.e., the molecular strains genuinely represent the same species), then the node (AB) representing a shared ancestor means little and does not contribute to cladogram resolution of evolutionary relationships between the species.

Past speciation events occurring on different strains force the strains apart on a molecular cladogram (Figs. 2-2, 3-1). Strains separated on a cladogram with descendant species in between produce phylogenetic paraphyly. When enough infraspecific strains are sampled, the paraphyly that appears may be interpreted as an ancestral species generating a descendant species. But if one or more molecular strains are extinct or unsampled, uncertainty in branch order is implicit. If paraphyly occurs at, say, one or two nodes distance between strains, then one might expect a similar level of paraphyly among all ancestral species, and given the widespread occurrence of paraphyly in published molecular cladograms, this is a major problem.

Random survival and extinction (or non-sampling) of molecular strains result in random branch order at the level of whatever paraphyly occurs in the group. Evaluation of morphological traits that are not phylogenetically informative can correct the branch order by identifying serial transformation of ancestral and descendant species.

Paraphyly can be simple, with branches for the same taxon only two nodes distant, or it can be more complex, with several nodes between branches leading to samples of the same taxon. Using "shared ancestry" as the sole methodology for phylogenetic reconstruction can result in highly precise but evolutionarily inaccurate (non-monophyletic) cladograms. Evaluation of serial transformations can correct this.

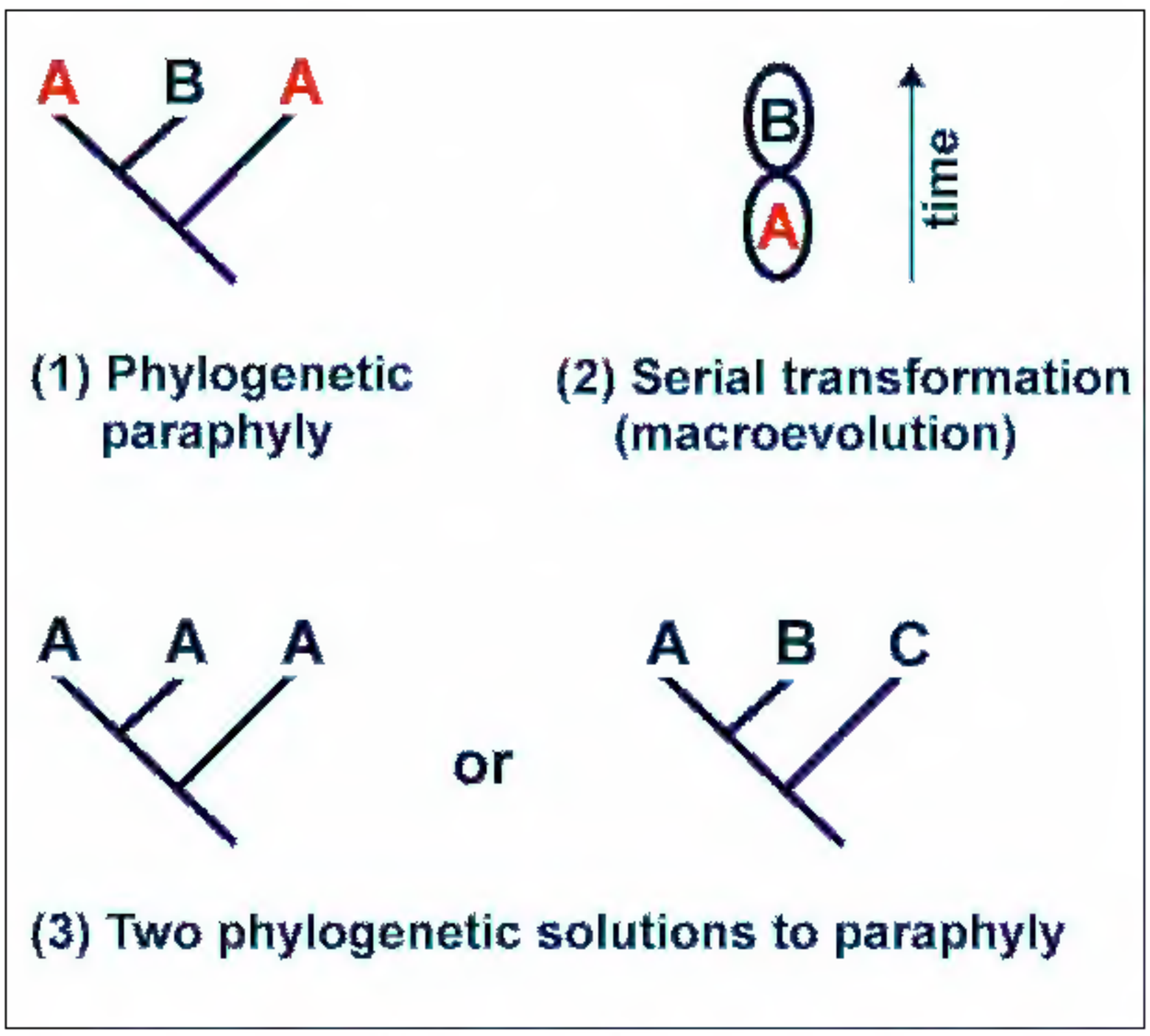

Figure 3. An example of simple phylogenetic paraphyly (1) is one taxon connected to two different nodes on a cladogram. Here species A is represented by two samples, these being distinct molecular strains that appear one node distant from each other (as in Figure 1). Evolutionary systematics suggests that if A is represented by two molecular strains, then B must arise from a deep ancestor of both samples of A. This is a serial transformation that may be represented by a simple caulogram (2). Phylogenetic systematics has two possible solutions (3) that do not abandon the cladogram, either lump the three samples into one taxon as on left, or have one of the A samples represent a new cryptic species as on right. Figure 3 shows problems in both morphological and molecular phylogenetic analysis, in fact at any rank, species through family.

Method

Bayesian statisticians who broke the Nazi war codes during World War II used small clues. The more clues, the better the analysis of the cryptographic code. The basic clue was one **deciban**, or **the minimum probability that will change belief in a hypothesis**. They used Bayesian statistical methods and the methods worked well. For instance, a change from a 0.50 probability (or 50:50 chance of being true) to about 0.55 in favor was considered the minimum clue that a particular hypothesis was true. As explained in detail in Part 3, decibans can be added together to get total probabilities that a hypothesis is true. This is because they are basically logarithmic. One deciban is 0.55 probability, five is 0.75, ten decibans is 0.90, 13 is 0.95, 20 is 0.99.

If you have ten clues of a minimum one deciban each that species A is ancestral to species B, then the chance is 0.90 that you are correct. If the traits examined as clues are very strong evidence of direction of evolution, then two traits at 7 decibans assigned to each will add to 14 decibans and then exceed 0.95 probability. Examples of trait transformations with high deciban values are diploidy → tetraploidy with no evidence of diploidization, widespread relictual habitats → recent habitats, sexual reproduction → asexual reproduction, and generalized morphology → unique specialization. These indications of macroevolutionary transformation between species are well documented in evolutionary texts as examples of (probably unidirectional) taxon transformations.

What you must do first, to make the analysis tractable, is to reduce if possible the species in the group you are studying to smaller groups with only one apparent generalist ancestor plus its apparent derivative species that radiate from it as transformations. These groups can be isolated from a standard dichotomous key to species, where derived species are commonly at the ends of series of couplets. These isolated groups will look like Fig. 1(3), or even Fig. 3(2), or as complex as Fig. 2 in Part 2. Criteria discussed in Part 2 of this paper are helpful. In some cases such groups may be treated as taxonomically distinct (the "dissilient genera" of Zander 2013: 92). Occasionally no generalist ancestor can be found among extant species for certain closely related apparently derived species, in which case a generalist shared ancestral species may be postulated.

For each group studied (in this paper, genera), the probabilities of one species transforming into another are calculated using decibans. Each species is assigned a number of clues that it is derived, based on theory of adaptive or at least rare or unique transformative trait transformations. The clues may be valued variously, but the restriction to 1, 3, 5, or 7 decibans controls guesswork and over-exactitude.

An expert can usually identify one species as apparent ancestor of a small or large cloud of derived species, at least in large genera. If there are more than one central ancestral species, segregation into separate genera (or other supraspecific rank) of each ancestor-species group is justified. A standard dichotomous key or a cladogram may be the first way to cluster species, or a "natural key" (Zander 2013: 47, 82) may be devised. Clues to which species are derived and which are ancestral are discussed in Part 2 of this paper. Assignment of deciban probabilities to various species clarifies direction and serial order of evolution.

For each group comprising one central ancestral species and attendant descendant species, if all species are clearly in that one group, and if one ancestor-descendant transformation is well supported, then all species are descendants of that identified ancestor. This allows one best probability to represent all of the transformation series. The probabilities may not be directly compared (they are not part of one probability distribution), however, because they are each generated by somewhat different data. They are, however, nested, so Bayes factors can be used. In this case Bayes factors are simply calculated as one probability (i.e., likelihood) for a particular hypothesis divided by the probability of another hypothesis. The significance of Bayes factors can be read off a standard table (see Part 2, Table 2).

A critical formula is the Implied Reliable Credible Interval (IRCI) calculation (see Part 2) which calculates the chance that at least one hypothesis has a high chance of being true in cases when you have a lot of hypotheses each with less than acceptable probability of being true. If the IRCI is 0.95 or greater, the question of which is ancestor is decidable. (If one species is demonstrably ancestor of one other species, then it is the ancestor of all of them in the group.) It is decidable is because there is sufficient information that at least one species is ancestor, but you do not know yet which one. A high Bayes factor can then identify which species is ancestor..

Some heuristics of monophyly

Heuristics are rules of thumb that help develop hypotheses or make decisions. Evolution involving modeling serial transformation of one taxon into another is complex and requires judgment, including evaluation of evidence. This paper formalizes (suggests a statistical basis for) analytic heuristics including the following:

1. Taxonomists use clues to evaluate monophyly. Sequential Bayes analysis through deciban assignment adds up clues in the context of a scale of probabilities.

2. Sequential Bayes analysis using decibans is a powerful analytic tool. It is similar to Shannon-Weaver information analysis, which is widely used in decision theory. There is a direct equivalence between decibans and bits, where 1 dB equals about 0.33 bits. The major relevant difference is that one deciban is considered the minimum for a perceived clue, but one bit (3 dB) is the minimum unequivocal unit of information (i.e., that which one may use to make a decision between two alternatives).

3. Identifying radiative transformation is a key element in monophyly estimation. Support for this comes from theory of adaptive radiation; Dollo non-reversal of large sets of traits, particularly those of whole taxa, being considered very rare in nature; and parsimony is enhanced by naming nodes in cladograms.

4. Direction of evolutionary change is clarified the more derived species there are. The IRCI formula shows how additional derived species contribute to increased probability of a particular species being the ancestor of all derived species.

5. The greater the polarization among species of a group, the more support there is for one particular species being ancestor. The more derived (advanced or different) species of a group are, the greater the Bayes factors are.

6. Extreme polarization can justify an unknown shared ancestor or a new genus. When certain derived species are more similar to each other than to the ancestor, either an unknown intermediate shared ancestor can be postulated to minimize the amount of difference from the ancestor, or a new genus can be proposed based on that support.

7. Transformational radiation in the context of monophyly provides support for recognition of taxa. The minimum of two well-supported (7 dB) traits to establish serial transformation of taxa parallels a standard heuristic of systematics that it takes at least two traits for a species to be well distinguished.

Short decision tree for analysis

1. If none of the species in a clear-cut group have high probability of being the ancestral taxon, and the IRCI formula does not indicate that there is a high probability that at least one taxon is the descendant of another; then the direction of transformation is undecidable based on present data. Else …

2. If none of the species in a clear-cut group have high probability of being the ancestral taxon, and the IRCI formula demonstrates a high probability that one species is ancestral to at least one other; then the direction of evolutionary transformation is decidable. Given that there is no other group for these species, they all are descendants of that one ancestral species even based on present somewhat skimpy data. The species with highest probability may be identified as ancestral if it has a high enough Bayes factor when compared with the species with second-highest probability of being the ancestor. Else …

3. If one of the species has a probability higher than 0.95 that it is ancestral to the others, and the other taxa have lower probabilities that they, themselves, are ancestral to others; then, these probabilities are converted to Bayes factors, which can compare probabilities of hypotheses calculated from somewhat different data sets. An acceptably high Bayes factor (consult Table 2 in Part 2) confirms that one species as ancestral to the others. This is the ideal scenario. Else …

4. If two species have probabilities higher than 0.95 that they are the ancestral taxon of the group, which can happen because the data sets are different; then, a tentative decision may be made based on how much higher the deciban count is for the best candidate as ancestral species over that of the second best. This is similar to reliance on the Akaiki information criterion which gauges the amount of information lost if the second-best alternative is chosen.

Contrived Examples

1. Probability of being ancestral for species A is 0.70, for species B is 0.50, for species C is 0.30. The IRCI formula (from the spreadsheet available online, see Part 2) indicates that the chance of at least one of the three being ancestor of one of the other two is 0.90, which is marginally undecidable. Bayes factor for A being ancestral to B is 1.4, which by the table in Part 2 is trivial. (This is calculated as likelihood of A, or 0.70, divided by probability of B, or 0.50, which is 7/5, or 1.4.)

2. Probability of being ancestral for species A is 0.70, for species B 0.50, for species C 0.40, and for species D, 0.40. The IRCI formula gives these four species a 0.95 chance that at least one of them is ancestral to at least one of the others. The hypothesis is then decidable. Species A is an acceptable candidate for ancestral species but the Bayes factor for species A giving rise to B is only 1.4. Both species remain good candidates for ancestral status.

3. Probability of being ancestral for species A is 0.90, for species B 0.25, for species C 0.20, and for species D, 0.20. The IRCI formula gives these four species a 0.95 chance that at least one of them is ancestral to at least one of the others. The hypothesis is then decidable. Species A is probably the ancestral species because the Bayes factor for species A giving rise to B is 3.2, or substantial.

4. Probability of species A is 0.95, of species B is 0.30, of species C is 0.30, of species D is 0.20. The IRCI formula is unnecessary because one species is probabilistically the ancestral species. The Bayes factor of species A versus species B is 3.2, which is substantial and it may be considered ancestral.

5. Probability of species A is 0.99, of species B is 0.95, of species C is 0.30, of species D is 0.20. The Bayes factor for A being ancestral to B is 1.02, which is trivial. However, the number of decibans represented by clues for a probability of 0.97 is 20, while for 0.95 it is 13. The weight of evidence in spite of the low Bayes factor is clearly in favor of species A being ancestral to species B and therefore to the remainder of the species in the group. The amount of information lost would be 7 dB if the species B were chosen as ancestor.

Remember that the probabilities do not have to add to 1.00 because somewhat different data sets are being used, so there is no probability distribution contributed by species C and D to weigh against species A being ancestral to species B.

Bayesian priors are often a cause for concern, because it is often thought that they are guesswork that may adversely affect the result. They are assumed in deciban analysis, however, to be equivocal, or 0.50. Only likelihoods based on empiric data are then used to calculate Bayes factors.

LITERATURE CITED

Zander, R.H. 2013. A Framework for Post-Phylogenetic Systematics. Zetetic Publications, St. Louis. CreateSpace Independent Publishing, Amazon.